DOCUMENT RESUME

ED 206 044 EA 013 761

AUTHOR Lysakowski, Richard S.; Walberg, Herbert J.
TITLE Classroom Reinforcement and Learning: A Quantitative Synthesis.
INSTITUTION Illinois Univ., Chicago. Chicago Circle Campus.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE 80
GRANT NIE-678-0090
NOTE 43p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 7-11, 1980). This document is Part G of a compilation entitled "Generalized Theory of Design and Analysis of Educational-Effectiveness Research." For related documents, see EA 013 759-760, EA 013 763-767, ED 168 714, ED 185 064, and EJ 210 984.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; Educational Research; Elementary Secondary Education; Learning Motivation; *Outcomes of Education; Postsecondary Education; Predictor Variables; *Reinforcement; Statistical Analysis

ABSTRACT

To estimate the influence of positive reinforcement on classroom learning, the authors analyzed statistical data from 39 studies spanning the years 1958-1978 and containing a combined sample of 4,842 students in 202 classes. Twenty-nine characteristics of each study's sample, methodology, and reliability were coded to measure their effects on reinforcement and learning. The types of reinforcement tested ranged from intangible social approval to tangible (even edible) objects. Statistical analysis using regressions and analysis of variance showed that reinforced students have markedly higher learning outcomes than do non-reinforced control students. Contrary to previous theory and opinions, the strong effects of instructional reinforcement appear constant across grades (kindergarten through college), socioeconomic levels, race, private and public schools, and community types. Tangible reinforcements may be slightly more effective than intangible ones, and female and special-school students may be slightly more affected by reinforcements, but generally there is little strong evidence that some types of reinforcers are more effective than others. (Author/RW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
[ ] Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED206044

Classroom Reinforcement and Learning: A Quantitative Synthesis

Richard S. Lysakowski and Herbert J. Walberg
University of Illinois at Chicago Circle

EA 013 761

Abstract

To estimate the instructional effects of rewards or positive reinforcement, both primary and secondary, on classroom learning, 102 effect sizes were calculated from statistical data in 39 studies spanning a period of 20 years and containing a combined sample of 4842 students in 202 classes. The mean of the study-median effect sizes is 1.17 which suggests average percentiles on learning outcomes of 88 and 50, respectively, for positively reinforced and control groups. Contrary to previous theory and opinions, the strong effects of instructional reinforcement appear constant across grades (kindergarten through college), socioeconomic levels, race, private and public schools, and community types; but tangible renforcers may be slightly more effective than intangible ones; and girls and students in special schools and classes may be slightly more affected by reinforcements than other students.

Classroom Reinforcement and Learning:
A Quantitative Synthesis

What aspects of rewards or reinforcement are most effective for learning? What techniques, on average, produce optimal results for different gender, race, and socioeconomic-status groups of students? The purpose of this paper is to provide a quantitative summary of recent research relevant to these and other questions by using statistical techniques developed by Glass (1978), Light and Smith (1971), Rosenthal (1976), and others to synthesize studies within a research domain. The present synthesis concerns experimental and quasi-experimental research in natural classroom settings of learning.

Theoretical Background

The idea that learning can be initiated, stimulated, or increased by conscious use of rewards is not a twentieth-century discovery. Aristotle and Plato argued that learners must be stimulated or motivated to awaken interest and desire for learning. Birnbaum (1962) describes how edible reinforcements (nuts, figs and honey) were used as rewards in the twelfth-century teaching of the Torah, (p. 385). Birnbaum (1975, p225) also cites medieval manuscripts, containing examples of social teachings, to parents and guardians, on the use of educational reinforcements, for example immediate and delayed reinforcements as well as social, tangible, and other types of reinforcements in the effort to induce children to acquire study habits and to love learning. In these writings, the developmental aspects of reinforcement are clearly described; and primary and secondary reinforcements such as food and praise are both exemplified.

Classic learning theorists of the current century, (Watson, 1930; Skinner, 1931; Thorndike, 1931; Dollard and Miller, 1950; Ferster and Skinner, 1957; Skinner, 1966) regarded reinforcements as essential components of successful

learning. Watson (1930) investigated the effects of reinforcements on changing behavior and attitudes towards learning. Skinner (1931) and E. L. Thorndike (1931) experimented at first on the effectiveness of various reinforcers on training animals, and subsequently transferred their interest and research activities to humans. Ferster and Skinner (1957) cooperated in producing exhaustive studies of learning responses to several schedules of reinforcers. Skinner (1966) later designed contingencies of reinforcement for an ideal culture.

Recent reviews extend the evidence on the general effectiveness of reinforcement in "token economies" and cooperative social structures. For example, O'Leary and Drabman (1971) found token-reinforcement programs generally effective in improving the academic and social behavior of diverse samples of children. Such programs are most effective, they found, when teacher, children, and parents are included in the planning stages and also in selecting systems of reinforcement and back-up reinforcers. Slavin (1977) and Michaels (1977) indicate that cooperative reward structures and group-oriented contingencies result in significantly strengthened cooperation among classroom groups as well as in individual improvement in learning.

A Reinforcement Hierarchy

Although recent reviews support the notion of the general efficacy of reinforcement, some theory and evidence suggest that reinforcers must be suited to the developmental, cultural, or social level of the learner. Forness (1973) for example, cites a number of experimental studies that suggest the developmental nature of reinforcement (Stevenson and Cruse, 1961; Zigler and Kanzer, 1962; Stevenson and Knight, 1961; Zigler and Butterfield, 1968). Teachers must gain knowledge of the learner's developmental stage, according to Forness, to know

what kind of reinforcement may be most appealing. Forness (1973) ordered reinforcements into seven developmental levels, from the most primitive such as edibles, bodily satisfactions, toys and money, through social acceptance on to the highest reinforcers such as competence or self-motivation.

Insert Table 1 about here

Forness (1973) based his categories or hierarchies broadly on psychological theory, but with the intention of making them relevant to classroom teaching practices.

Table 1 shows Forness' (1973) hierarchy of instructional reinforcers. The first three categories (competence, feedback and social approval) according to Forness are those reinforcers which exist naturally in the regular classrooms. The last four categories are those that are more often administered in special programs for children with problems in functioning in regular classes.

Forness repeats the ancient suggestion that there is a developmental aspect of the reinforcer hierarchy. From conclusions reached by several experimenters (Stevenson & Cruse, 1961; Zigler & Butterfield, 1968), it appears to Forness (1973) and Havighurst (1970) that the primary techniques are most effective with the youngest children and children in the lowest developmental, social, or cultural stages. As children develop and expand their intellectual capacity, they should be "promoted", suggests Forness, (1973) to higher categories of reinforcers. There comes an age, or a developmental stage, wherein students can no longer be influenced or motivated by tokens such as paper (silver-or gold-colored) stars, candies, or even material trinkets. Then social approval or verbal praise is more effective than material rewards or even money (Gregory, 1976; Cradler, 1971). Finally, the intrinsic or

"competence" level is reached by those students who are intellectually most developed, and most gifted, for whom only awareness of being right is sufficient reinforcement, as they feel compelled by the pleasure of solving a problem or the sheer satisfaction of acquiring a skill. (Forness (1973) remarks that it is doubtful that in the present-day American educational system many students get to that point. In the present synthesis, only one experiment was found (Baron, 1972) to have conditions that appear to correspond to the definition of this category).

Adoption of any of the seven categories of reinforcements, according to Forness (1973) following Havighurst (1970), should be based on the stage of development of the personal characteristics of the learner such as age level, physical and intellectual level, The student's race and socioeconomic status, in Forness' view, usually have a bearing on what type of reward would be most effective. Forness also mentions the need for different kinds of reinforcers for students in special classes, for whom rewards offered in ordinary class settings are not appealing enough.

Havighurst (1970) related his reward hierarchy to cultural, social and socio-economic status of the learners. Havighurst (1970) argues for the importance of understanding sub-cultural attitudes regarding what constitutes a personal success. The motivational rewards must then be chosen so as to represent the closest approach to the ideals of the given sub-culture. According to Havighurst (1970) as well as Forness (1973) reinforcers must be matched to students' developmental level and other personal characteristics for maximum effectiveness.

Schultz and Sherman (1976) criticize the developmental theory proposed by Havighurst (1970) and Forness (1973) on grounds of insufficient evidence to prove that social class-by-reinforcer interactions yield consistent results. Schultz and Sherman, however, conducted no systematic search and analysis of evidence on this point. In addition to the primary purpose of the present paper, to synthesize the overall results of reinforcements in classrooms, the evidence on differential impacts of reinforcers is examined.

Method

Literature Search and Selection

Four sources were used in the search for articles and studies dealing with educational reinforcements:

1. Original references (starting points of the search) were found in the writings of Block (1977), Bloom (1976), Medley (1978), and Rosenshine (1971) which review many of the studies on the effectiveness of teaching methods for much of the pre-1974 period.

2. Forward references, that is, published works citing the references in 1.) were in the Social Sciences Citation Index (SSCI) for the 1975 through 1978 issues.

3. The Educational Resources Information Center (ERIC) provided several of the current references.

4. The Bibliographical Retrieval System, (BRS), a computer-assisted service, yielded other recent references. From this system of information retrieval, two data bases were used, the ERIC system and the Psychological Abstracts International, both for the period covering years 1975-1978, inclusive.

In all sources, studies and reports were obtained by using the following descriptors: Academic Achievement, Reinforcement, Reinforcers Motivation Feedback, with Preschool, Kindergarten, Elementary Level, Primary Grades, Secondary School, Secondary Education in High Schools, Quality Instruction, Cues, Behavior Change, Behavior Modification, Participation, Participation-involvement, and Rewards with Achievement. All references were screened for relevance to the subject of classroom reinforcement.

The following questions were asked in the selection process:

1. Does the report or article have any relevance to the subject of instructional reinforcemet?

2. Does the source contain any quantitative information or data?

3. Is the data available in the form needed to conform with the design of this synthesis, i.e., can the effect size be computed by any of the several methods possible such as the means and standard deviation of control and the experimental groups or pre- and post-treatment conditions, (Glass, 1976)?

4. If the means and standard deviation are not available, are the individual scores, for calculating the means and standard deviations or other statistics available?

Having used the above search methods and selection criteria, the counts of the studies obtained through the BRS computer search were as follows: the ERIC system yielded 40 studies, of which 22 were examined closely and only 6 were selected; Psychological Abstracts yielded 98 studies, of which 43 were examined and 7 selected. Thus the total number of studies selected and used in the quantitative analysis from the two data bases of the computer-assisted search was 13. Of the other 26 studies used in the

analysis, only one (Clark and Walberg, 1968) was taken from Bloom's (1976) original review since it reported means and standard deviations; but 25 studies came from "branch" references, i. e., references found in studies located through one of the four sources listed previously, even though those studies themselves could not be used.

Of the 138 studies identified, only 39 could be used in the quantitative synthesis, since the others lack sufficient quantitative data. An example of the types of studies omitted from quantitative analysis here is Walker and Buckley's (1972), a study of the effects of reinforcement, punishment and feedback on academic response rate. This study, though it presented interesting aspects of the effects of punishment on overall academic response, did not include in the report, individual scores, standard deviations, or T-tests; and the effect size could not be computed. Another example of a study omitted from the analysis was Clemens and Tracy (1977). In this study the effects of tactile and verbal reinforcement on the classroom behavior and arithmetic performance of emotionally disturbed boys were analyzed. Significant achievements effects were claimed; but only the medians of the group rates were reported, which did not correspond to the model criteria set in the present synthesis, (in this case individual scores for both the experimental and control groups would have been sufficient). Anecdotal data and interesting theories are mentioned in these and similar omitted studies; and the writers have no reason to think that the studies left out would change the results dramatically if somehow the original data could be retrieved.

Two special populations, namely students who are mentally retarded and learning-disabled, were excluded from the search. These special populations

belong in a class by themselves, and limitations of time and space of this synthesis dictated the imposition of these restrictions. College and adult students have not been specifically searched for, but reports of experiments of such groups encountered during the search, which related to the topic of classroom reinforcement, have been included (for example, Komaki, 1975).

Listings of 75 unpublished doctoral dissertations and other unpublished work were not searched for the purpose of this synthesis. The search for these unpublished sources now would, in the writers' opinion, be excessively costly and would unduly extend the bounds of this paper, without contributing commensurately to the clarification of the principles presented here.

Of the studies which had quantitative data reported, the information sought was such that could be adapted for calculating the "effect size" as suggested by Glass (1976). For a detailed description of the "effect size" and methods of computing it, see the Coding and Analysis on procedure.

It seems likely that the bulk of the usable published studies on the subject of educational reinforcement, within the time period 1958-1978, have been found and incorporated in the present analysis. Some researchers, however, may have used titles that were not descriptive enough or did not fall within the range of criteria for the search and selection process.

Coding and Analysis

Recent writings of Gage (1978), Glass and Smith (1977), Light and Smith (1971), and Rosenthal (1976) make a good case for the need and value of quantitative synthesis in educational and psychological research. A new technique of quantitative synthesis, otherwise referred to as meta-analysis in educational and psychological research was convincingly presented by Glass (1978).

The present research findings are analyzed and synthesized in compliance with the techniques suggested by Glass (1978).

In order to make many diverse experimental methods and procedures comparable, meta-analysis converts the results of all experiments to an "effect size" which may be thought of as a measure of effectiveness of a particular treatment given to an experimental group of learners versus the results from a conventional or control group.

The general form of the effect size is defined as:

$$ES = \frac{\bar{X}_e - \bar{X}_c}{S_c}$$

where $\bar{x}_e$ represents the mean of the experimental group's results, $\bar{X}_c$ = the mean of the control group, and S_c = the Std. Deviation of the control group. There are also formulas (Glass, 1976) for calculating ESs from F-values or t-values on correlations are given, plus the numbers of cases for both groups. Thus, a study to be usable in the present design of synthesis had to show some of these few essential values to enable the writers to calculate the effect size.

For the purpose of the present synthesis, twenty-nine characteristics were coded for each of the 102 effect sizes from 39 studies selected. The characteristics coded are explained in Table 2, but several points require

Insert Table 2 about here

additional comment. Since some studies grouped categories such as boys and girls together in reporting results, "mixed" categories were used on some factors. The first author coded all studies, and the other coded ten independently. Since there was more than 95 percent agreement, the second author

checked the other 29 studies. In personal communications, Gene V Glass (1978) Edward Haertel (1979), and Kent Bailey (1979) provided advice on the calculating effect sizes in complex cases of factorial studies and repeated measures, although the authors assume responsibility for all errors.

In the analyses of variance, all 29 factors were employed as nominal variables, grouped and indicated as in Table 2. In the regressions, however, the continuous variables, age (grade level) and sample size, were left in their full metric precision without grouping; and the nominal factors were recoded to sets of binary (0,1) variables to assess the possible effect of the value of each category on the effect sizes.

Because the number of comparisons varies among studies, studies with greater numbers of comparisons would be weighted more heavily than others if each comparison were given equal weight. To give each study equal weight in the analysis, each comparison was given a weight inversly proportional to the total number of comparisons in the study from which it was taken; for example, Marholin et al (1974), twelve comparisons each received a weight of 1/12.

Even though weighted analyses weight each study equally, they do not remove statistical dependencies among the comparisons within each study that violate inferential assumptions. The jackknife procedure (Tukey, 1977) was used on selected variables to provide stringent, independent estimates of the regression coefficients and their standard errors.

Results and Discussion

The characteristics of students in the studies researched covered a broad spectrum. In the age variable, there were 236 preschoolers, 2198 in elementary schools, 2293 in secondary and 295 in college. Three races were distinguished individually in the studies: 454 black, 831 white and 145 oriental students. The majority of the studies do not mention the racial make up of their student group, even though it is quite likely that other races may have been included in the groups studied. Only two studies distinguished girls separately, which included a total of 32 girls, and twenty-eight studies showed separate data for boys, comprising 540 boys; the rest of the comparisons did not distinguish the sexes and were treated as "mixed".

Table 3 shows a frequency distribution for the study-median effect

Insert Table 3 about here

sizes and all 102 effect sizes as well as the percentages within each class interval. For the study-median effect sizes, the mean effect-size is 1.2 and the standard deviation is 1.0; the median is .9 and the range extends from - 0.1 to +8.5.

Table 3 shows two extreme upper effect sizes, 8.5 and 8.1 far from the main cluster of the effect sizes. Both these outlying results come from studies of special-education classes (O'Leary and Becker, 1967; and Solomon and Wahler, 1973) and deal with elementary level students with a history of abnormally disruptive behavior. In one case (Solomon and Wahler, 1973), peer reinforcement was used as a technique of behavior modification.

In the second study, (O'Leary and Becker, 1967), edibles and trinkets were used as token reinforcement items to obtain highly significant effect sizes of over 8.0. Such atypical cases would unduly influence the results of the analyses. To offset the atypical effect of the outliers, the upper extreme two studies, 8.5 and 8.1, were "Winsorized" for the purpose of the regression analysis, i.e., their values were reduced to the next highest value in Table 3, 3.8 (Dixon and Massey, 1969, pp. 328-332).

Table 4 shows the results of analysis of variance for the 29 factors

Insert Table 4 about here

for all 102 effect sizes and also for study-weighted sizes. The more significant borderline factors in weighted analysis are sample size, gender, school type, contamination and compensation, and peer-group control.

To control the factors for one another, a series of simple and multiple regressions were run with the effect size on all other variables separately. From these regressions it became apparent which variables had the highest and statistically significant correlation with learning effects.

Next, 15 variables were selected that were significant at .1, .05 and .01 levels. Table 5 shows the results of five main effect size regressions. Column 1 represents the simple regressions of the effect size against the individual variables. The simple regressions indicate that only the special schools and female students are statistically significant (probabilities less than .05.)

Column 2 of Table 5 shows the multiple regressions for the 15 independent variables. In this analysis, only the special school variable is significant. Columns 4 and 5 show that the exclusion of the most powerful variable, special

schools, from the regression model does not appreciably change the significance levels of the remaining variables. Column 6 in Table 5 shows the final jackknifed regression which can be considered robust against the undue influence of a single study on the regression statistics. The regression shows that girls (at the .10 level) and students in special schools (at the .05 level) are more influenced by reinforcement in general than are other students. None of the other variables such as student characteristics, reinforcement type, and methodological validity of the study appear to affect the generally large impact of reinforcement.

Table 6 shows that adding cross-products of sample characteristics with the seven types of reinforcement raises the accountable variance in the effect sizes very slightly and insignificantly, which means that the various types of reinforcement have no detectably different effects on females, blacks, lower socio-economic groups or children in the lower elementary grades or in small samples than on other students, even though reinforcement in general may have greater effects on girls and students in special schools.

In conclusion, the general effect of reinforcement in classrooms is moderately large and fairly robust. The mean effect-size of 1.2 suggests that experimentally-reinforced groups in the sample scored on average at about the 88th percentile in comparison to 50th percentile for the control groups. It appears that the differences in reinforced and control groups are higher in samples of students in special schools and, possibly, of girls. There is no significant evidence, however, that some types of reinforcers are more effective than others nor differentially effective for younger, female, black, or lower socio-economic samples.

Table 1

Definitions of Reinforcement Constructs

1. Competence (autonomy): denotes also learning for learning's sake;
2. Feedback (being correct): being told only the results of the test taken, or task performed; it implies a sense of competition in this hierarchy;
3. Social approval (non-contingent praise): ranges from a glance or a slight smile to effusive social praise; non-competitive;
4. Contingent activity: follows the principle wherein the teacher uses a high-frequency behavior, such as watching an enjoyable T.V. music lesson, to reinforce a low-frequency activity, e. g. attaining a 90% correct level in a math test (Madsen et al., 1976); the teacher tells students about the contingency beforehand;
5. Tokens (or check marks); tokens are accumulated for a number of steps in an on-going task and then exchanged for the most desired reinforcer (tangible or intangible);
6. Tangibles (non-edible): the reinforcers are given to students directly upon completion of a task; the reinforcers (toys or trinkets) are displayed before the task is begun;
7. Edibles: the rules are exactly the same as in 6., above , except the reinforcers are edible (cookies, candies etc.) which the students are asked to consume.

Table 2

Information coded for Each Study

	Coding, Comment or Sample Items
Identification	
Author	
year of study	
number of study	
number of comparison within study	
Sample characteristics	
Sex	boys = 1; girls = 2; mixed = 3
Grade level	preschool = 0; first grade = 1; second grade = 2; etc; college freshmen = 13; sophmores = 14; juniors = 15; seniors = 16; adults other than college = 17.
Subject's race	black = 1; white = 2; latino = 3; oriental = 4; mixed = 5; other ethnic = 6.
Experimenter's race	black = 1; white = 2
SES	low = 1; middle = 2; upper = 3
Region	east = 1; middle = 2; midwest = 3 west = 4; south = 5; other country = 5
School type	public = 1; private = 2; special education = 3 (mental health center, special education class in a regular school, or other institutional school).
Community type	urban = 1; rural = 2; suburban = 3; other = 4.

Table 2 continued

Methodological characteristics	
Control technique	stratification = 1; partial correlation = 2; means (not a control) = 3; beta or standard weights in regression = 4; raw or metric weights in regression = 5; factorial analysis of variance = 6; analysis of covariance = 7.
Educational outcomes	academic achievement, factual knowledge = 1; academic achievement, higher level (comprehension, understanding, concept attainment) = 2 I. Q. scores = 3; grades 4; discrimination learning = 5; accuracy-mathematics = 6; attempted-mathematics = 7; behavior modifications = 8; personality inventory = 9; accuracy of task performance = 10; attempted task performance = 11; accuracy-English = 12; attempted-English = 13; retention = 14.
Reinforcement techniques	competence = 1; being correct = 2; see Table 4 for a complete description of reinforcement categories.
Ability	inadequate or unspecified = 1; adequate = 2; an indication that S's pre-treatment ability was measured or ascertained;
Home environment	omitted or inadequate indication = 1; adequate = 2.

Table 2 continued

Motivation	inadequate or unspecified = 1; adequate = 2.
Peer group influence	inadequate or unspecified = 1; adequate = 2.
Classroom environment	inadequately measured or unspecified = 1; adequate = 2.
<u>Threats to validity</u>	
Reliability of measures	Are the outcome measurement instruments internally consistent or stable enough to detect changes during, or differences in, treatments? inadequate or unspecified = 1; adequate = 2 (.75 or greater).
Reliability of treatment implementation.	Have the units in a treatment actually undergone the same conditions? inadequate or unspecified = 1; adequate = 2 (evidence shown in the report).
Maturation	Have factors within the units, rather than the presumed treatment causes, brought about changes or differences? (Extended treatment or without random assignment to treatments and control conditions) inadequate or unspecified = 1; adequate = 2 (evidence inferred from the report).

Table 2 continued

Selection bias	Did pre-existing differences among the groups account for later differences on the dependent variable? (Evidence of uncontrolled pre-existing differences). inadequate or unspecified = 1; adequate = 2 (evidence shown in the report).
Contamination and compensation	Did untreated control groups imitate or somehow gain the benefits of treatment; or work harder to compete with the experimental group? (Evidence of teacher exchange or other bias); inadequate or unspecified = 1; adequate = 2 (evidence shown in the report)
Mortality	Did differential drop-out rates in the treatment, or control groups account for differences on the dependent variable obtained at the end of the experiment? (self evident) inadequate or unspecified = 1;
Generalizability	Can the cause - effect be generalized to other times, units, and settings, classes or schools? inadequate or unspecified = 1; adequate = 2 (evidence shown in the report)
Operational irrelevancy	Did the treatment administered contain Hawthorne effect, experimenter bias, placebo or other unintended causes that produced the effects? inadequate or unspecified = 1; adequate = 2 (self-evident in the report).

Table 2 continued

Statistical power	Was the sample large enough to reject the null hypothesis at a given level of probability or were estimate coefficients with a reasonably small margin of error? (60 or more classes, 100 or more individuals)? inadequate or unspecified = 1, adequate = 2

Table 3

Frequency Distribution

of Effect Sizes

Class Intervals	Effect sizes Frequency	Effect sizes Percentage	Study-Median Effect sizes Frequency	Study-Median Effect sizes Percentage
9.0-9.4				
8.5-8.9	1	1.0	1	2.6
8.0-8.4	1	1.0	1	2.6
.				
.				
.				
4.0-4.4				
3.5-3.9	2	2.0		
3.0-3.4	3	2.9	2	5.1
2.5-2.9	4	3.9		
2.0-2.4	7	6.9	2	5.1
1.5-1.9	7	6.9	4	10.3
1.0-1.4	20	19.5	6	15.4
.5-.9	37	36.2	17	43.5
.0-.4	13	12.8	6	15.4
-.0- - .4	5	4.9		
-.5 - -.9	1	1.0		
-1.0 - - 1.4				
-1.5- - 1.9	1	1.0		
-2.0- - 2.4				
Total	102	100.0	39	100.0

Table 4

Analysis of Variance Statistics

For Each Factor

Factor and Level Variable	Comparison-Weighted Analysis					Study-Weighted Analysis				
	$\bar{X}$	S	N	F	p.	$\bar{X}$	S	N	F	p.
Sample Size				5.09	.003				3.58	.023
N = 2-49	1.41	1.13	51			1.70	1.19	17		
N = 50-82	.83	.81	24			.84	.78	11		
N = 83-150	.89	.36	15			.91	.47	3		
N = 151-2139	.40	.57	12			.47	.50	7		
Gender				5.43	.006				2.40	.105
Boys	.75	1.00	28			.95	1.01	9		
Girls	2.84	.44	2			2.99	----	1		
Mixed	1.16	.93	72			1.15	1.01	28		
Grade Level				.28	.843				.03	.994
Preschool	.84	.84	6			.98	.88	2		
Elementary	1.12	1.00	82			1.18	1.07	28		
Secondary	.89	1.11	10			1.17	1.37	5		
College	1.10	.53	4			1.20	.58	3		
Race of the Students				1.66	.165				.10	.983
Black	1.46	1.09	15			1.27	1.04	7		
White	.70	.97	25			1.04	1.26	9		
Oriental	1.34	.06	2			1.34	----	1		
Mixed	.89	.15	3			.88	.15	2		
Unspecified	1.15	.097	57			1.21	1.07	20		

Factor and Level Variable	Comparison-Weighted Analysis					Study-Weighted Analysis				
	$\bar{X}$	S	N	F	p.	$\bar{X}$	S	N	F	p.
Experimenter Race				1.49	.226				1.11	.299
Black	1.12	1.05	87			1.23	1.08	35		
White	.79	.33	15			.65	.31	4		
Socio-economic Status				.25	.777				.47	.630
Lower	1.16	1.14	22			1.36	1.22	8		
Middle	1.12	.95	44			.97	1.01	15		
Unspecified	.99	.95	36			1.27	1.01	15		
Region				5.36	.0002				.88	.507
East	1.65	1.57	4			1.56	1.14	6		
Midwest	.66	1.19	12			.88	.93	13		
West	.85	.83	37			.82	.66	8		
South	1.12	.51	22			1.35	1.16	7		
Other Country	1.90	.92	13			1.82	----	2		
Unspecified	1.47	1.57	4			1.69	1.84	3		
School Type				6.07	.001				8.22	.0003
Public	.79	.58	60			.79	.55	26		
Private	1.60	1.09	16			1.30	1.22	3		
Special Classes	1.51	1.40	24			2.34	1.31	9		
Unspecified	.20	.08	2			.20	----	1		
Community Type				2.09	.107				.40	.757
Urban	.97	.87	52			1.08	.93	21		
Rural	.83	1.01	21			1.21	1.30	7		
Suburb	1.46	1.13	16			.93	1.03	4		
Unspecified	1.44	1.08	13			1.54	1.27	7		

Table 4 continued

Factor and Level Variable	Comparison-Weighted Analysis					Study-Weighted Analysis				
	$\bar{X}$	S	N	F	p.	$\bar{X}$	S	N	F	p.
Control Technique				3.65	.003				.63	.682
Stratification	1.44	1.03	9			1.68	1.19	6		
Partial Correlation	1.06	.95	3			1.01	.96	2		
Beta Weights	.70	.90	15			.97	----	2		
Raw Weights	1.99	1.06	10			1.99	----	1		
ANOVA	.97	.91	56			1.00	1.00	23		
Anal. of covariance	.39	.31	6			.47	.30	2		
Unspecified	2.13	.91	3			2.38	1.10	2		
Educational Outcome				2.06	.028				.69	.694
Acad. Achievement, factual knowledge	1.33	.92	33			1.30	1.01	8		
Acad. Achievement, higher level; comprehension & concept attainment	1.22	.47	5			1.33	.55	3		
I.Q. tests	.48	.51	3			.48	---	1		
Grades	1.12	.99	4			.82	.0	2		
Discrimination Learning	.44	.53	4			.45	.46	3		
Accuracy-Math	.72	.14	3			---	---	---		
Attempted-Math	.91	.71	4			1.29	---	1		
Behavior Modification	1.09	1.10	26			1.44	1.24	13		
Personality inventory	2.17	1.51	5			1.91	0.0	2		
Accuracy-Task	.70	.59	5			.48	---	1		
Attempted-Task	.76	.54	6			.54	.49	2		
Accuracy-English	.55	.11	2			---	---	---		
Attempted-English	-.93	.90	2			---	---	---		

Table 4 continued

Factor and Level Variable	Comparison Weighted Analysis					Study-Weighted Analysis				
	$\bar{X}$	S	N	F	p.	$\bar{X}$	S	N	F	p.
Instrument Type				5.03	.027				.19	.665
Local	.84	.97	45			1.10	1.12	21		
Standard	1.27	.96	57			1.25	.90	18		
Instructional Reinforcers				1.15	.314				.26	.99
Monetary	.96	.40	2			----	---	---		
Social	.52	.55	6			.44	.40	4		
Verbal praise and tactile	1.24	1.16	14			1.17	1.26	4		
Chart	.55	----	1			----	----	---		
Instrinsic	.52	----	1			----	----	---		
Written notes	.60	.49	3			.52	----	1		
Token, edible	1.47	1.61	8			1.87	1.76	4		
Token, intangible	2.11	1.28	4			1.86	----	1		
Criticism	.27	.62	5			.35	.74	2		
Punishment	1.98	.48	2			1.98	----	1		
Grades	.98	----	1			.98	----	1		
Chart and Monetary	2.45	----	1			----	----	---		
Token, tangible and response cost	1.65	1.08	2			----	----	---		
Token, intangible and response cost	1.45	1.05	2			----	----	---		
No tokens, tutoring	2.14	.44	3			----	----	---		
T. V. music lessons	1.19	----	1			----	----	---		
Immediate reinforcement	.92	.32	4			.94	----	1		
Delayed reinforcement	.79	.84	2			----	----	---		

Factor and Level Variable	Comparison Weighted Analysis					Study-Weighted Analysis				
	X	S	N	F	p.	X	S	N	F	p.
Non-contingent reinforcement	.14	----	1			----	----	---		
End-of-session reinforcement	.05	----	1			----	----	---		
Material (candy)	.81	.18	4			----	----	---		
Symbolic (e.g.+mark)	1.68	----	1			----	----	---		
Extrinsic rewards and punishment	.94	.32	4			.94	.39	2		
Group contingencies	1.00	.74	3			1.00	----	1		
Assistant (T-aide)	2.33	.24	2			----	----	---		
Correctness	.52	.18	3			----	----	---		
Token, peer tutor	2.10	2.41	2			2.66	----	1		
Vicarious	2.21	----	1			2.21	----	1		
Contingency (reading)	.82	.65	8			1.21	.74	2		
Contingency (chance reinforcement)	.40	1.07	6			----	----	---		
Contingency (Math fun activity)	.90	.21	2			.90	----	1		
No-award vs. exp. award	.73	----	1			----	----	---		
Unexp. vs, exp. award.	.77	----	1			----	----	---		
Ability				.22	.639				.64	.428
Adequately controlled	1.12	.95	53			1.29	1.08	22		
Omitted or inadequate	1.03	1.03	49			1.02	1.00	17		
Motivation				2.34	1.30				.54	.466
Adequately controlled	.93	.89	52			1.05	.99	20		
Omitted or inadequate	1.23	1.06	50			1.30	1.11	19		

Table 4 continued

Factor and Level Variable	Comparison Weighted Analysis					Study-Weighted Analysis				
	X	S	N	F	p.	X	S	N	F	p.
Home environment				1.69	.20				.86	.359
Adequately controlled	1.46	1.39	10			1.63	1.39	4		
Omitted or inadequate	1.04	.93	92			1.12	1.01	35		
Peer group				2.65	.11				2.93	.096
Adequately controlled	1.43	1.37	17			1.68	1.43	9		
Omitted or inadequate	1.01	.88	85			1.02	.87	30		
Class environment				2.86	.09				2.45	.126
Adequately controlled	1.41	1.31	82			1.58	1.38	11		
Omitted or inadequate	1.00	.88	20			1.01	.86	28		
Reliability of measures				1.53	.220				2.46	.125
Adequate	1.27	1.01	29			1.49	1.16	15		
Omitted or inadequate	1.00	.97	73			.96	.94	24		
Reliability of treatment				.05	.816				.00	.980
Adequately controlled	1.08	1.00	97			1.17	1.07	36		
Omitted or inadequate	.98	.72	5			1.18	.92	3		
Maturation				.05	.828				.30	.587
Adequately controlled	1.08	1.00	96			1.20	1.07	36		
Omitted or inadequate	.99	.76	6			.85	.74	3		
History				.05	.828				.30	.587
Adequately controlled	1.08	1.00	96			1.20	1.07	36		
Omitted or inadequate	.99	.76	6			.85	.74	3		
Selection bias				.45	.505				.50	.484
Adequately controlled	1.06	.98	93			1.12	1.03	33		
Omitted or inadequate	1.29	1.09	9			1.46	1.22	5		

Table 4 continued

Factor and Level Variable	Comparison Weighted Analysis					Study-Weighted Analysis				
	X	S	N	F	p.	X	S	N	F	p.
Contamination and compensation				2.41	.124				3.00	.091
Adequately controlled	1.04	.96	95			1.07	.97	35		
Omitted or inadequate	1.63	1.24	7			2.00	1.44	4		
Mortality				.06	.809				.33	.568
Adequately controlled	1.09	1.01	95			1.20	1.08	35		
Omitted or inadequate	.99	.70	7			.88	.61	4		
Generalizability				2.94	.090				2.29	.139
Adequately controlled	.94	.90	61			.99	.97	25		
Omitted or inadequate	1.28	1.08	41			1.51	1.12	13		
Operational irrelevance				.59	.557				.27	.767
Adequately controlled	1.06	1.01	88			1.15	1.09	32		
Omitted or inadequate	1.21	.68	14			1.18	.87	7		
Statistical power				3.10	.046				2.07	.141
Adequate	.72	.55	28			.66	.56	10		
Omitted or inadequate	1.22	1.06	74			1.33	1.14	28		
Reinforcement hierarchy				1.97	.077				1.39	.252
Competence	.52	----	1			----	----	---		
Being correct	.72	.52	7			.57	.61	2		
Social approval	1.14	1.00	32			1.06	.93	15		
Contingent activity	.78	.74	35			.88	.61	11		
Tokens	1.72	1.33	15			1.88	1.55	5		
Tangibles	1.24	.71	5			1.25	----	1		
Edibles	1.24	1.21	7			2.21	2.09	2		

Table 5

Five Main Effect Regression Models

	1		2		3		4		5		6	
Variables	Bivariate		Multivariate		Select Multivariate		Select Multivariate		Select Multivariate		Jackknifed Sel. Multivariate	
	b	T	b	T	b	T	b	T	b	T	b	T
Sample Size	-.001	1.406	-.000	.712			-.003					
Mall	-.288	.703	-.521	1.410			-.236					
Female	1.887	2.145	.937	.931			1.930	2.098	1.994	2.350	1.208	1.6
Private Schools	.138	.217	.121	.190								
Special Schools	1.522	4.834	1.204	2.703	1.522	4.834					1.731	2.2
Unspecified School Type	-.990	.934	-.855	.908								
Peer Group	.662	1.711	.060	.109								
Class Environment	.571	1.565	.104	.212			.299					
Contamination & Compensation	-.931	1.733	-.622	1.110			-.994	1.765	-1.007	1.982	-.630	.1
Reliability of Measures	.529	1.569	.243	.696			.415					
Social Approval	-.176	.510	----	----			.192					
Contingent Activity	-.403	1.101	-.091	.263			.091					

G 30

Table 5

(Continued)

	1		2		3		4		5		6	
Variables	Bivariate		Multivariate		Select Multivariate		Select Multivariate		Select Multivariate		Jackknifed Sel. Multivariate	
	b	T	b	T	b	T	b	T	b	T	b	T
Tokens	.832	1.823	---	---			.457					
Tangible Reinforcement	.000	.000	.263	.330			.321					
Edible Reinforcement	1.102	1.536	.604	.882			1.126					
Constant			1.453		.817		1.592		2.004		2.153	
Multiple R			.743		.622		.631				.672	

Table 6

Multiple Correlations for Five

Controlled Stepwise Regressions

	Main Effect Model	Main Effects and Product Model
Female	.659	.678
Black	.677	.696
Socioeconomic Status	.680	.689
Grade Level	.676	.684
Sample Size	.667	.735

Appendix

Allen, S. A., Spear, P. S., & Lucke, J. R. Effects of social reinforcement on learning and retention in children. Developmental Psychology. 1971, 5, 73-80.

Allen, S. The effects of verbal reinforcement on children's performance as a function of type of task. Journal of Experimental Child Psychology. 1966, 3, 57-73.

Baron, R. M., and Ganz, R. L. Effects of locus of control and type of feedback on the task performance of lower-class black children. Journal of Personality and Social Psychology. 1972, 21, 124-130.

Brown, D., Reschley, D., & Sabers, D. Using group contingencies with punishment and positive reinforcement to modify aggressive behavior in a head start classroom. The Psychological Record. 1974, 24, 491-496.

Brown, R. A., & Epstein, J. Interaction of achievement level and reinforcing properties of daily grading system. Education. 1977, 98, 131-134.

Buckholdt, D. R., and Wodarski, J. S. The effects of different reinforcement systems on cooperative behaviors exhibited by children in classroom contexts. August, 1974. Paper presented at the annual meeting of the American Psychological Association, New Orleans, Louisiana.

Carringer, D. and Wilson, C.S. The effects of sex, socioeconomic class, experimenter race, and kind of verbal reinforcement on the performance of black children. The Journal of Negro Education. 1974, 43, 212-220

Clark, M., Lachowicz, J., & Wolf, M. A pilot basic education program for school dropouts incorporating a token reinforcement system. Behavior Research & Therapy. 1968, 6, 183-188.

Clark, C. A., & Walberg, H. J. The influence of massive rewards on reading achievement in potential urban school dropouts. American Educational Research Journal. 1968, 5, 305-311.

Clements, J. E., and Tracy, D. B. Effect of punitive dress on classroom behavior of institutionalized emotionally disturbed adolescents. Psychological Reports. 1976, 39, 335-338.

Clingman, J. M., Auerback, S. M., Bowman, P. C., and Parrish, J. M. Differential effects of candy, social, and token rewards on the I. Q. scores of children of above average intelligence. Psychology in the schools. 1977, 14, 95-98.

Cradler, J. D., and Goodwin, D. L. Conditioning of verbal behavior as a function of age, social class, and type of reinforcement. Journal of Educational Psychology. 1971, 62, 279-284.

Epstein, M. H., Repp, A. C. and Cullinan, D. Decreasing "obscene" language of behaviorally disordered children through the use of a DRL schedule. Psychology in the Schools. 1978, 15, 419-423.

Evans, D. R., Horvath, P., Sanders, S., and Dolan, J. Reinforcement of attention and academic performance in a special education class. Psychological Reports. 1974, 35, 1143-1146.

Frederiksen, L. W., and Frederiksen, C. B. Teacher-determined and self-determined token reinforcement in a special education classroom. Behavior Therapy, 1975, 6, 310-314.

Gregory, M. K. Effects of locus of control and type of reinforcement of adolescent boys. The Journal of Educational Research. 1976, 70, 45-49.

Henderson, W. T., and Shih-Sung Wen. Effects of immediate positive reinforcement on undergraduates' course achievement. Psychology Reports. 1976, 39, 568-570.

Johnson, R. A. Differential effects of reward versus no-reward instructions on the creative thinking of two economic levels of elementary school children. Journal of Educational Psychology. 1974, 66, 530-533.

Kazdin, A. E. Vicarious reinforcement and direction of behavior change in the classroom. Behavior Therapy. 1977, 8, 57-63.

Kennelly, K., and Kinley, S. Perceived contingency of teacher administered reinforcements and academic performance of boys. Psychology in the Schools. 1975, 12, 449-453.

Klein, R. D., Thompson, M. The use of an individualized motivation system to modify accuracy and rate on a computer-based number facts program. Paper presented at the Annual Meeting of the AERA (San Francisco, Cal., April, 19-23, 1976).

Komaki, J. Neglected reinforcers in the college classroom. Journal of Higher Education. 1975, 46, 63-73.

Lepper, M. R., Green, D., and Nisbett, R. E. Undermining children's intrinsic interest with extrinsic reward: A test of the "Overjustification Hypothesis". Journal of Personality and Social Psychology. 1973, 28, 129-137.

Madsen, C. K., Dorow, L. G., Moore, R. S. and Womble, J. U. Effect of music via television as reinforcement for correct mathematics. Journal of Research in Music Education. 1976, 24, 51-59.

Marholin, D., McInnis, E. T., and Heads, T. B. Effect of two freetime reinforcement procedures on academic performance in a class of behavior problem children. Journal of Educational Psychology. 1974, 66, 872-879.

Miller, L. B., and Estes, B. W. Monetary reward and motivation in discrimination learning. Journal of Experimental Psychology. 1961, 61, 501-504.

O'Leary, K. D., and Becker, W. C. Behavior modification of an adjustment class: A token reinforcement program. Exceptional Children, 1967, May, 637-642.

Page, D. P. and Edwards, R. P. Behavior change strategies for reducing disruptive classroom behavior. Psychology in the Schools. 1978, 15, 413-418.

Page, E. B. Teacher comments and student performance: A seventy-four classroom experiment in school motivation. The Journal of Educational Psychology. 1958, 49, 173-181.

Rosenfeld, G. W. Some effects of reinforcement on achievement and behavior in a regular classroom. Journal of Educational Psychology. 1972, 63, 189-193.

Smeets, P. M., and Striefel S. The effects of different reinforcement conditions on the test performance of multihandicapped deaf children. Journal of Applied Behavior Analysis. 1975, 8, 83-89.

Solomon, R. W. and Wahler, R. G. Peer reinforcement control of classroom problem behavior. Journal of Applied Behavior Analysis, 1973, 6, 49-56.

Sorensen, R. L. Attainment value and type of reinforcement: A hypothesized interaction effect. Journal of Personality and Social Psychology, 1976, 34, 1155-1160.

Spear, P. S. and Spear, S. A. Social reinforcement, discrimination learning and retention in children. Developmental Psychology. 1972, 7, 220.

Spear, P. S. and Armstrong, S. Effects of performance expectancies created by peer comparison as related to social reinforcement, task difficulty, and age of child. Journal of Experimental Child Psychology, 1978, 25, 254-266.

Taffel, S. J., and O'Leary, K. D. Reinforcing math with more math: choosing special academic activities as a reward for academic performance. Journal of Educational Psychology, 1976, 68, 579-587.

Tichenor, J. L. Self-monitoring and self-reinforcement of studying by college students. Psychological Reports. 1977, 40, 103-108.

Washington, K. R. The effects of systematic reinforcement and a self-awareness program on the self-concept of black preschool children. Child Study Journal, 1976, 6, 199-208.

Winefield, Helen. Remedial reading: effects of tokens, material back-up reinforcers, and response cost. Perceptual and Motor Skills, 1977, 44, 383-394.

References

Birnbaum, P. A treasury of Judaism. New York: Hebrew Publishing Co., 1962.

Birnbaum, P. A book of Jewish Concepts. New York: Hebrew Publishing Co., 1975.

Block, J. H., & Tierney, M. L. An exploration of two correction procedures used in Mastery Learning approaches to instruction. Journal of Educational Psychology. 1974, 66, 962-967.

Block, J. H. (Ed.) Schools, Society and Mastery Learning. New York: Holt, Rinehart and Winston, Inc., 1974.

Bloom, B. S. Human characteristics and school learning. New York: McGraw-Hill Book Co., 1976.

Dixon, W. J., and Massey F. J., Jr. Introduction to statistical analysis. New York: McGraw-Hill Book Co., 1969.

Dollard, J. and Miller, N.E. Personality and psychotherapy. New York: McGraw-Hill, 1950.

Ferster, C. B. Arbitrary and natural reinforcement. Psychological Record. 1967, 17, 341-347.

Forness, S. R. The reinforcement hierarchy. Psychology of the Schools, 1973, 10, 168-177.

Gage, N. L. The scientific basis of the art of teaching. New York: Teachers College Press, 1978.

Glass, G. V. Primary, secondary and meta-analysis of research. Educational Research, 1976, 5, 3-8.

Glass, G. V. Integrating findings: the meta-analysis of research. Review of Research in Education, 1976, 5, 351-379.

Havighurst, J. Minority subcultures and the law of effect. The American Psychologist. 1970, 25, 313-322.

Hilgard, E. R., and Bower, G. H. Theories of learning. New York: Appleton-Century-Crofts. Educational Division of Meredith Corporation, 1971.

Light, R. J., and Smith, P. V. Accumulating evidence: Procedures for resolving contradictions among different research studies; Harvard Educational Review, 1971, 4, 429-471.

MacMillan, D., & Forness, S. Behavior modification: savior or savant? American Journal of Mental Deficiency, in press.

Medley, D. M. The effectiveness of teachers. In Peterson, P. L., and Walberg, H. J., editors, Research on Teaching; Concepts, Findings and Implications. Berkeley, Cal., McCutchan Publishing Corporation, 1979.

O'Leary, K. D., and Drabman, R. Token reinforcement programs in the classroom: A review. Psychological Bulletin, 1971, 75, 379-398.

Rosenshine, B. V. Content, time and direct instruction. In Peterson, P. L., and Walberg, H. J., editors, Research on Teaching; Concepts, Findings, and Implications. Berkeley, Cal.: McCutchan Publishing Corporation, 1979.

Schultz, B., and Sherman, R. H. Social class, development, and differences in reinforcer effectiveness. Review of Educational Research, 1976, 46, 25-59.

Skinner, B. F. Contingencies of reinforcement in the design of a culture. Behavioral Science, 1966, 11, 159-166.

Skinner, B. F. Contingencies of Reinforcment, A Theoretical Analysis. New York: Appleton-Century-Cross, Div. of Meredith Corporation, 1969.

Stevenson, H. W., & Cruse, D. B. The effectiveness of social reinforcement with normal and feeble-minded children. Journal of Personality, 1961, 29, 124-135.

Thorndike, E. L. Human learning. New York: The Century Co., 1968.

Thorndike, Edward L. Animal intelligence, experimental studies. New York: Hafner Publishing Co., Inc. 1911. Reprinted 1970.

Tukey, J. W. Exploratory data analysis. Reading, Mass.: Addison-Wesley, 1977.

Watson, J. B. Behaviorism. Chicago: The University of Chicago Press; 1930.

Ziegler, E., & Butterfield, E. Motivational aspects of changes in IQ test performance of culturally deprived nursery school children. Child Development, 1968, 39, 1-14.